自回归移动平均模型分析

本章介绍自回归移动平均模型分析，包括平稳时间序列的自回归移动平均模型分析的思路、模型识别、参数估计、检验和预测，以及非平稳时间序列的自回归求积移动平均模型分析。

第一节 自回归移动平均模型

第二节 自回归移动平均模型的识别

第三节 自回归移动平均模型的估计

第四节 **ARMA**模型检验和预测

第一节 自回归移动平均模型

一、时间序列分析的BJ方法论

- Box和Jenkins在《时间序列分析：预测和控制》中提出，时间序列的本质特征就是相邻观测值的依赖性，大部分平稳时间序列可以用一系列自回归移动平均随机过程加以模拟。
- 这引出了时间序列分析的一种重要思路——通过时间序列数据背后生成数据的自回归移动平均过程模型进行分析。
- 这种时间序列分析方法被称为"BJ方法论"，属于时域分析方法，是现代时间序列计量经济学最重要的方法论之一。

- 时间序列分析BJ方法论的理论基础

时间序列数据可以看作由一些离散型随机过程生成的，是特定离散型随机过程的一次实现，时间序列数据的性质由生成它们的随机过程决定，因此可以通过识别和分析时间序列背后的随机过程模型，进行时间序列数据的分析和预测。

- 运用BJ方法论进行应用分析包括四个步骤：

1、**识别**

根据各种时间序列模型的理论特征，对时间序列数据进行分析，确定适当的初步模型，包括模型类型及其阶数，就是找出ARMA模型适当的p、q值；

2、**估计**

用适当的参数估计方法，估计初步设定模型的相关参数值，包括自回归和移动平均系数和白噪声的方差等；

3、诊断

对模型进行校验，包括检验模型的拟合程度，检验设定模型的合理性和阶数是否正确等，并进一步调整修改模型，确定模型和精确估计参数；

4、预测和控制

利用所得到的模型进行预测分析，包括静态预测和动态预测，多步预测等，利用模型进行控制。预测本身也是对模型的进一步检验。

二、自回归移动平均模型

（一）移动平均模型（moving average process, MA）

- 移动平均过程就是一个白噪声过程不同时间随机变量的加权和。
- 最简单的移动平均过程是当期和前一期白噪声的加权和：

$$Y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}$$

这种移动平均过程称为"一阶移动平均过程"，记为MA(1)。

- 较复杂的移动平均过程可以包括多个甚至无限个滞后项的加权和，例如：

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} = \mu + \sum_{j=0}^{q} \theta_j \varepsilon_{t-j}$$

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} = \mu + \sum_{j=0}^{q} \theta_j \varepsilon_{t-j}$$

- 这些模型分别记为MA(2)、MA(*q*)和MA(∞)。

引进滞后算子L（$L\varepsilon_t=\varepsilon_{t-1}, L^2\varepsilon_t=\varepsilon_{t-2}, \cdots$ ），移动平均模型可分别表示为： Y_t

$$Y_t=\mu+\varepsilon_t+\theta\varepsilon_{t-1}=\mu+(1+\theta L)\varepsilon_t=\mu+\varphi(L)\varepsilon_t$$

$$Y_t=\mu+\varepsilon_t+\theta_1\varepsilon_{t-1}+\theta_2\varepsilon_{t-2}=\mu+(1+\theta_1 L+\theta_2 L^2)\varepsilon_t=\mu+\varphi_2(L)\varepsilon_t$$

$$Y_t=\mu+\sum_{j=0}^{q}\theta_j\varepsilon_{t-j}=\mu+(1+\theta_1 L+\theta_2 L^2+\cdots+\theta_q L^q)\varepsilon_t=\mu+\varphi_q(L)\varepsilon_t$$

$$Y_t=\mu+\sum_{j=0}^{\infty}\theta_j\varepsilon_{t-j}=\mu+(1+\theta_1 L+\theta_2 L^2+\cdots)\varepsilon_t=\mu+\varphi_\infty(L)\varepsilon_t$$

（二）自回归模型（autoregressive process，AR）

- 时间序列的当前值取决于其自身若干期或无限期前期水平。
- 自回归过程同样有一阶、二阶、阶或无穷阶自回归模型，分别记为AR(1)、AR(2)、AR()和AR(∞)，表达式分别为

$$Y_t = \eta + \phi Y_{t-1} + \varepsilon_t$$

$$Y_t = \eta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$$

$$Y_t = \eta + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t = \eta + \sum_{i=1}^{p} \phi_i Y_{t-i} + \varepsilon_t$$

$$Y_t = \eta + \phi_1 Y_{t-1} + \cdots + \varepsilon_t = \eta + \sum_{i=1}^{\infty} \phi_i Y_{t-i} + \varepsilon_t$$

- 引进滞后算子表示方法，上述AR模型则可以分别表示为：

$$(1-\phi L)Y_t = \Phi_1(L)Y_t = \eta + \varepsilon_t$$

$$(1-\phi_1 L-\phi_2 L^2)Y_t = \Phi_2(L)Y_t = \eta + \varepsilon_t$$

$$(1-\phi_1 L-\phi_2 L^2-\cdots-\phi_p L^p)Y_t = \Phi_p(L)Y_t = \eta + \varepsilon_t$$

$$(1-\phi_1 L-\phi_2 L^2-\cdots)Y_t = \Phi_\infty(L)Y_t = \eta + \varepsilon_t$$

（三）自回归滑动平均模型

- 既包含一系列白噪声扰动的加权平均，也包含时间序列本身滞后项加权平均的混合时间序列过程。这样的过程称为"自回归移动平均过程"，记为ARMA。
- 最简单的自回归移动平均过程包括一阶自回归项和一阶移动平均项，即

$$Y_t = \eta + \phi_1 Y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

记为ARMA(1,1)。

- 一般的自回归移动平均过程包括*p*阶自回归和*q*阶移动平均，即

$$Y_t = \eta + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

也可以写成

$$Y_t = \eta + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{i=0}^{q} \theta_j \varepsilon_{t-j}$$

其中 $\theta_0 = 1$ ，记为ARMA(*p*,*q*)。

（四）自回归求积移动平均模型（autoregressive integrated moving average process, ARIMA）

- 对于非平稳的时间序列，可以先利用单积（差分）方法把它们转化成平稳序列，然后再利用平稳时间序列的ARMA模型进行分析。
- 一个经单位根检验确定为非平稳的时间序列，运用单积分析判定经d次差分后变为平稳的，那么该经过d次差分得到的序列就可进一步用ARMA(p,q)中某一种模型来表示。这时原时间序列被记为ARIMA(p,d,q)，其中d是单积阶数，p是自回归阶数，q是移动平均阶数。

三、ARMA模型平稳和可逆的条件

（一）移动平均模型的平稳和可逆条件

1、MA(1)模型

（1）平稳性条件

根据MA(1)模型的定义和白噪声序列的性质，很容易知道对MA(1)模型满足：

$$E(Y_t)=E(\mu+\varepsilon_t+\theta\varepsilon_{t-1})=\mu$$

$$\gamma_0=Var(Y_t)=E(Y_t-\mu)^2=E(\varepsilon_t+\theta\varepsilon_{t-1})^2=\sigma_\varepsilon^2+\theta^2\sigma_\varepsilon^2=\sigma_\varepsilon^2(1+\theta^2)$$

$$\gamma_k=Cov(Y_t,Y_{t-k})=E[(Y_t-\mu)(Y_{t-k}-\mu)]=\begin{cases}\theta\sigma_\varepsilon^2 & k=1\\ 0 & k>1\end{cases}$$

- MA(1)的均值、方差和 $k=1$ 和 $k>1$ 时自协方差都是不随时间变化的常数。根据时间序列的平稳性可知，MA(1)模型是协方差平稳的。

（2）可逆性

- 移动平均过程的可逆性指移动平均过程的扰动，可以反过来用时间序列的当前和过去水平表示，从而移动平均过程可以转化为自回归过程。
- 把MA(1)模型 $Y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}$ 改写成：

$\widetilde{Y}_t = Y_t - \mu = \varepsilon_t + \theta\varepsilon_{t-1}$ 或 $\varepsilon_t = \widetilde{Y}_t - \theta\varepsilon_{t-1}$

- 通过迭代可得到

$$\varepsilon_t = \tilde{Y}_t - \theta(\tilde{Y}_{t-1} - \theta\varepsilon_{t-2}) = \tilde{Y}_t - \theta\tilde{Y}_{t-1} + \theta^2\varepsilon_{t-2}$$

$$= \cdots = \tilde{Y}_t + \sum_{j=1}^{\infty}(-\theta)^j Y_{t-j}$$

也可以写成： $$Y_t = \varepsilon_t - \sum_{j=1}^{\infty}(-\theta)^j Y_{t-j}$$

只有当 $|\theta| < 1$时，上述MA(1)模型才能把扰动表示成时间序列当前和过去水平的形式，或者把MA(1)模型表示成自回归形式，也就是MA(1)模型才是**可逆**的。

2、MA(*q*)模型

（1）平稳性

- 根据MA(*q*)的定义得到：

$$E(Y_t)=\mu$$

$$\gamma_0=Var(Y_t)=\sigma_\varepsilon^2(1+\theta_1^2+\cdots+\theta_q^2)$$

$$\gamma_k=Cov(Y_t,Y_{t-k})=\begin{cases}(\theta_k+\theta_{k+1}\theta_1+\theta_{k+2}\theta_2+\cdots+\theta_q\theta_{q-k})\sigma_\varepsilon^2 & k=1,\cdots,q\\ 0 & k>q\end{cases}$$

（2）可逆性

- MA(q)模型 $Y_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \cdots + \theta_q\varepsilon_{t-q}$

可以用滞后算子写成：

$$Y_t - \mu = (1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q)\varepsilon_t$$

由于只有当

$$1 + \theta_1 Z + \theta_2 Z^2 + \cdots + \theta_q Z^q = 0$$

的根都落在单位圆之外，确保可以根据滞后算子多项式的逆运算得到：

$$\varepsilon_t = (1+\theta_1 L+\theta_2 L^2+\cdots+\theta_q L^q)^{-1}(Y_t-\mu)$$

$$= (1+\eta_1 L+\eta_2 L^2+\eta_3 L^3+\cdots)(Y_t-\mu)$$

- 也就是把MA(*q*)模型表示成了AR(∞)形式，因此上述滞后算子多项式方程的根都在单位圆之外，就是MA(*q*)模型可逆性的条件。

3、无限项移动平均过程MA(∞)

- 这种时间序列过程不是无条件平稳的。
- 可以证明当系数序列满足平方可加性($\sum_{j=0}^{\infty}\theta_j^2 < \infty$)或绝对可加性($\sum_{j=0}^{\infty}\left|\theta_j\right| < \infty$)（后者强于前者，隐含前者）时，MA(∞)是协方差平稳的。
- 平方可加性和绝对可加性的实质是，间隔时间较长的扰动的影响会很快下降，很快趋向于0。

- 当系数序列绝对可加时，MA(∞)过程的均值、方差和协方差，都可以从MA(∞)的相应结果简单推广得到：

$$E(Y_t)=\lim_{j\to\infty}E(\mu+\varepsilon_t+\theta_1\varepsilon_{t-1}+\theta_2\varepsilon_{t-2}+\cdots+\theta_j\varepsilon_{t-j})=\mu$$

$$\gamma_0=Var(Y_t)=\lim_{j\to\infty}E(\varepsilon_t+\theta_1\varepsilon_{t-1}+\theta_2\varepsilon_{t-2}+\cdots+\theta_j\varepsilon_{t-j})^2$$

$$=\lim_{j\to\infty}\sigma_\varepsilon^2(1+\theta_1^2+\cdots+\theta_j^2)$$

$$\gamma_k=Cov(Y_t,Y_{t-k})=(\theta_k+\theta_{k+1}\theta_1+\theta_{k+2}\theta_2+\cdots)\sigma_\varepsilon^2$$

- 当移动平均模型参数满足绝对（或平方）可加性时，上述期望和方差、协方差都是与时间无关的有限数。
- 因此对于移动平均过程来说，只要阶数有限则肯定平稳，若阶数无限必须满足上述绝对（或平方）可加性时才平稳。
- 一个平稳时间序列过程不可能由不满足上述系数条件的无限移动平均过程产生。

[例] 讨论下列移动平均过程模型的平稳性和可逆性：

$$Y_t = \mu + \varepsilon_t + 0.6\varepsilon_{t-1} - 0.2\varepsilon_{t-2}$$

- 首先因为有限阶移动平均模型都是平稳的，因此该模型肯定是平稳的。
- 其次，该模型的特征多项式为

$$1 + 0.6Z - 0.2Z^2 = 0$$

两个根分别为4.19和1.19，都在单位圆之外，因此也是可逆的。

（二）自回归模型的平稳性条件

1、AR(1)模型

- $|\phi| \geq 1$时，AR(1)不是协方差平稳的。
- AR(1)模型通过迭代可以转换成了一个MA(∞)模型：

$$Y_t = \eta + \phi Y_{t-1} + \varepsilon_t = \eta + \varepsilon_t + \phi Y_{t-1} = (\eta + \varepsilon_t) + \phi(\eta + \varepsilon_{t-1} + \phi Y_{t-2})$$

$$= (\eta + \varepsilon_t) + \phi(\eta + \varepsilon_{t-1}) + \phi^2(\eta + \varepsilon_{t-2}) + \cdots$$

$$= \frac{\eta}{1-\phi} + \varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \cdots$$

若令 $\frac{\eta}{1-\phi}=\mu$，$\theta_j=\phi^j$，则AR(1)模型转化为：

$$Y_t=\mu+\varepsilon_t+\theta_1\varepsilon_{t-1}+\theta_2\varepsilon_{t-2}+\cdots=\mu+\sum_{j=0}^{\infty}\theta_j\varepsilon_{t-j}$$

与MA(∞)模型完全一致。

- 根据MA(∞)模型的平稳性条件，不难知道AR(1)模型平稳的条件 ϕ^j 为有绝对可加性。当 $|\phi|=1$ 或 $|\phi|>1$ 时，绝对可加性都不成立，因此 $|\phi|<1$ 是AR(1)模型平稳的条件。